

ABSTRAK

Curah hujan dengan intensitas yang bervariasi dapat memberikan dampak signifikan terhadap berbagai sektor kehidupan, khususnya di wilayah perkotaan seperti Kota Surabaya yang memiliki pola curah hujan yang cukup fluktuatif. Variabilitas tersebut menjadikan prediksi curah hujan penting untuk mendukung berbagai aktivitas dan pengambilan keputusan yang berkaitan dengan kondisi cuaca. Dalam beberapa tahun terakhir, pendekatan *machine learning* banyak dimanfaatkan untuk meningkatkan akurasi prediksi fenomena meteorologis. Salah satu algoritma yang banyak digunakan adalah *Extreme Gradient Boosting* (XGBoost), yaitu algoritma berbasis *ensemble boosting* yang mampu menghasilkan performa prediksi tinggi serta efektif dalam menangani hubungan nonlinier antarvariabel meteorologis. Namun demikian, model XGBoost cenderung sulit diinterpretasikan sehingga diperlukan metode interpretasi model seperti *SHapley Additive exPlanations* (SHAP) untuk menjelaskan kontribusi setiap variabel terhadap hasil prediksi. Penelitian ini bertujuan membangun model klasifikasi curah hujan di Kota Surabaya menggunakan XGBoost dengan interpretasi SHAP. Data yang digunakan berupa data meteorologis harian Stasiun Meteorologi Maritim Perak I BMKG periode Januari 2018 hingga September 2025 dengan variabel suhu, kelembapan, lama penyinaran matahari, serta arah dan kecepatan angin. Variabel target dikategorikan menjadi empat kelas, yaitu tidak hujan, hujan ringan, hujan sedang, serta hujan lebat hingga sangat lebat. Data dibagi dengan rasio 80:20 serta dilakukan *hyperparameter tuning* menggunakan *GridSearchCV* dan *stratified k-fold cross-validation* ($k = 5$). *Synthetic Minority Oversampling Technique* (SMOTE) diterapkan pada data latih karena distribusi data tidak seimbang. Hasil penelitian menunjukkan bahwa model XGBoost mampu mengklasifikasikan kategori curah hujan dengan akurasi sebesar 80,04%, sementara analisis SHAP menunjukkan bahwa suhu minimum merupakan variabel paling berpengaruh terhadap prediksi hujan sedang hingga hujan lebat di Kota Surabaya. Pendekatan ini tidak hanya menghasilkan model yang akurat tetapi juga memberikan interpretasi terhadap faktor meteorologis yang memengaruhi curah hujan.

Kata kunci: Curah hujan, XGBoost, SHAP, Klasifikasi, BMKG.